ФИЗИКА ЭЛЕМЕНТАРНЫХ ЧАСТИЦ И АТОМНОГО ЯДРА
2025. Т. 56, вып. 3. С. 1175–1183

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВЫЧИСЛЕНИЯ ФУНКЦИОНАЛЬНЫХ ИНТЕГРАЛОВ В КВАНТОВОЙ ТЕОРИИ ПОЛЯ

А. В. Васильев [1], *А. С. Иванов* [1], *Д. В. Сальников* [1,2], *В. В. Чистяков* [1,*]

[1] Московский государственный университет им. М. В. Ломоносова, Москва
[2] Институт ядерных исследований РАН, Москва

Рассматривается проблема вычисления функциональных интегралов в моделях квантовой механики. С помощью нейросетевого алгоритма нормализующего потока вычислены функциональные интегралы в моделях гармонического осциллятора, потенциала двойной ямы и потенциала Морзе. Выполнено сравнение с результатами моделирования методом Монте-Карло.

The problem of computing functional integrals in quantum mechanics models is considered. Functional integrals in models of the harmonic oscillator, double well potential, and Morse potential were computed using a neural network algorithm for normalizing flows. A comparison was made with results obtained from Monte Carlo simulation.

PACS: 03.65.−w; 02.70.Ss

ВВЕДЕНИЕ

Поставим задачу вычисления $\langle 0|F|0\rangle$ — среднего значения от квантово-механической наблюдаемой F по основному состоянию гамильтониана:

$$H = \frac{p^2}{2m} + V(x). \tag{1}$$

Среднее по основному состоянию равно, как известно, среднему по распределению Гиббса в пределе низких температур:

$$\langle F\rangle_\beta = \frac{\sum\limits_n \mathrm{e}^{-\beta(E_n-E_0)}\,\langle n|F|n\rangle}{\sum\limits_n \mathrm{e}^{-\beta(E_n-E_0)}}, \quad \lim_{\beta\to\infty}\langle F\rangle_\beta = \langle 0|F|0\rangle. \tag{2}$$

* E-mail: vsevolod.chistyakov@gmail.com

С другой стороны, среднее по распределению Гиббса представляется в виде функционального интеграла с периодическими граничными условиями $x(0) = x(\beta)$ [1]:

$$\langle F \rangle_\beta = \frac{1}{Z} \int D[x(t)] \exp\left(- \int\limits_0^\beta dt \left[\frac{m\dot{x}^2}{2} + V(x) \right] \right) F(x(t)). \tag{3}$$

При произвольном $V(x)$ интеграл (3) не вычисляется аналитически. В настоящей работе мы вычисляем интеграл (3) численно, применяя генеративные алгоритмы машинного обучения. Введем на отрезке $[0, \beta]$ решетку с шагом τ, число узлов решетки $N = \beta/\tau$. Функции $x(t)$ поставим в соответствие вектор $x \in \mathbb{R}^N$, компонентами которого являются значения функции в узлах решетки.

Интеграл (3) аппроксимируется интегралом конечной кратности

$$\langle F \rangle_\beta = \int d^N x P(x) F(x) + O(\tau^2), \quad x \in \mathbb{R}^N, \tag{4}$$

$$\begin{gathered} P(x) = \frac{\mathrm{e}^{-S(x)}}{Z}, \\ S(x) = \tau \sum_{n=1}^{N} \left(\frac{m(x_n - x_{n-1})^2}{2\tau^2} + V(x_n) \right). \end{gathered} \tag{5}$$

Вычислим интеграл (4) методом Монте-Карло. Для этого необходимо сгенерировать набор N-мерных векторов, $\{x^{(k)}\}$, распределенных с плотностью вероятности $P(x)$ [2]. Интеграл (4) приближенно равен среднему по набору

$$\int d^N x P(x) F(x) \approx \frac{1}{M} \sum_{k=1}^{M} F(x^{(k)}). \tag{6}$$

Задача вычисления какой-либо наблюдаемой F сводится, таким образом, к задаче генерации многомерных векторов (траекторий) с распределением $P(x)$. Стандартный подход к решению этой задачи заключается в применении алгоритма Метрополиса — строится марковская цепь, финальным распределением которой является $P(x)$. Такой алгоритм требует больших временны́х затрат. Существует подход [3–6], в котором для ускорения генерации траекторий применяется нейросетевой генеративный алгоритм нормализующих потоков. В настоящей работе мы используем этот алгоритм в различных моделях квантовой механики, в том числе тестируем его на точно решаемых моделях.

1. ГАРМОНИЧЕСКИЙ ОСЦИЛЛЯТОР НА РЕШЕТКЕ. АНАЛИТИЧЕСКИЕ СООТНОШЕНИЯ

Рассмотрим гармонический осциллятор. Выберем систему единиц, в которой $\hbar = m = \omega = 1$, ω — частота осциллятора. Действие (5) представимо в виде квадратичной формы

$$S(x) = \frac{1}{2}\,(Ax, x), \quad A = \frac{1}{\tau}\left((2+\tau^2)\, I - T - T^\dagger\right),$$

где T — образующая группы $\mathbb{Z}_N$ (в регулярном представлении):

$$T_{i,k} = \delta_{i+1,k}, \quad T^\dagger = T^{-1}, \quad T^N = I.$$

Построим ортогональную матрицу O из следующих столбцов:

$$(u_k)^s = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi k s}{N}\right), \quad (w_k)^s = \sqrt{\frac{2}{N}} \sin\left(\frac{\pi k s}{N}\right), \qquad (7)$$
$$O = ||u_0, u_1, w_1, \ldots, u_{N/2-1}, w_{N/2-1}, u_{N/2}||.$$

Преобразование O приводит T к блочно-диагональному виду (регулярное представление $\mathbb{Z}_N$ раскладывается на сумму неприводимых):

$$O^\dagger T O = 1 \bigoplus R_1 \bigoplus \ldots \bigoplus R_{N/2-1} \bigoplus (-1),$$

где R_k — матрица поворота на угол $(2\pi k)/N$.

Оно также приводит матрицу A к диагональному виду: $O^\dagger A O = \mathrm{diag}\,(\Lambda_0, \Lambda_1, \Lambda_1, \ldots)$.

Матрица $C = On$, где $n = \mathrm{diag}\,(\Lambda_0^{-1/2}, \Lambda_1^{-1/2}, \ldots)$, преобразует матрицу A к единичной:

$$C^T A C = I, \quad A^{-1} = C C^T.$$

Задача генерации траекторий в случае гармонического осциллятора имеет аналитическое решение. Пусть имеется набор N-мерных векторов $z \sim \mathcal{N}^N(0,1)$, координаты которых распределены нормально со средним 0 и дисперсией 1. Тогда векторы $x = Cz$ будут иметь распределение $P(x) = \exp\left((-1/2)(Ax, x)\right)/Z$.

2. НОРМАЛИЗУЮЩИЙ ПОТОК

Если потенциал $V(x)$ имеет более сложный вид, преобразование $x = g(z)$, отображающее нормально распределенный набор векторов $z \sim \mathcal{N}^N(0,1)$ в набор векторов $x \sim P(x)$, будет нелинейным и в общем случае не выражается аналитически. Однако можно поставить задачу аппроксимации искомого отображения функцией из некоторого многопараметрического семейства.

Будем искать функцию g в виде композиции аффинных преобразований [7]

$$g = \mathcal{A}_n \circ \ldots \circ \mathcal{A}_{i+1} \circ \mathcal{A}_i \circ \ldots \circ \mathcal{A}_1,$$

где каждое преобразование $\mathcal{A}_i$ действует следующим образом. Представим вектор z в виде прямой суммы векторов одинаковой длины $z = u \oplus v$

$$\mathcal{A}(u) = u, \quad [\mathcal{A}(v)]_k = \mathrm{e}^{\theta_{1k}(u)}\, v_k + \theta_{2k}(u),$$

где $\theta : \mathbb{R}^{N/2} \to \mathbb{R}^N$ — некоторая функция, принадлежащая многопараметрическому семейству функций: $\theta(u) \equiv \theta(u|w)$, где $w \in \mathbb{R}^K$ — параметры.

Функция θ имеет следующий вид: $\theta = \sigma(w_m \sigma(\ldots w_1 u)))$, где w_k — произвольные матрицы, σ — нелинейное преобразование (функция активации). Таким образом, θ представляет собой полносвязную нейросеть с m скрытых слоев.

С увеличением размерности генерируемых объектов необходимо увеличивать также и количество параметров генеративной модели [4]: количество аффинных преобразований, скрытых слоев и размер скрытого слоя. В настоящей работе, следуя [6], мы обучаем генеративную модель, в которой количество аффинных преобразований $n = 16$, число скрытых слоев $m = 12$, размер скрытого слоя $\sim N$. Дальнейшее увеличение числа обучаемых параметров не приводит к заметному улучшению качества генерации. Используется многомасштабная архитектура [5].

Для обучения нейронной сети (подбора параметров w) минимизируем функцию потерь [5, 8]

$$L[w] = \frac{1}{M} \sum_{k=1}^{M} \left(S(g(z^{(k)}|w)) - \ln \left| \det \frac{\partial g(z^{(k)}|w)}{\partial z} \right| \right).$$

3. ГАРМОНИЧЕСКИЙ ОСЦИЛЛЯТОР. ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Двухточечная функция Грина $G_s = (1/N) \sum\limits_{i=1}^{N} x_i\, x_{i+s}$ и квадрат волновой функции, вычисленные по траекториям, сгенерированным нейросетью, с большой точностью приближают теоретические значения величин (рис. 1).

Если после преобразования $g(z)$ применить также ортогональное преобразование матрицей (7), то таким образом возможно учесть в генеративной модели трансляционную $\mathbb{Z}_N$ симметрию задачи. Это существенно ускорит сходимость нейросети. Мы обучали две генеративные модели в течение одной эпохи, в одном случае применялось ортогональное преобразование, а в другом нет. В первом случае качество генерации уже после одной эпохи обучения достаточно высокое (рис. 2).

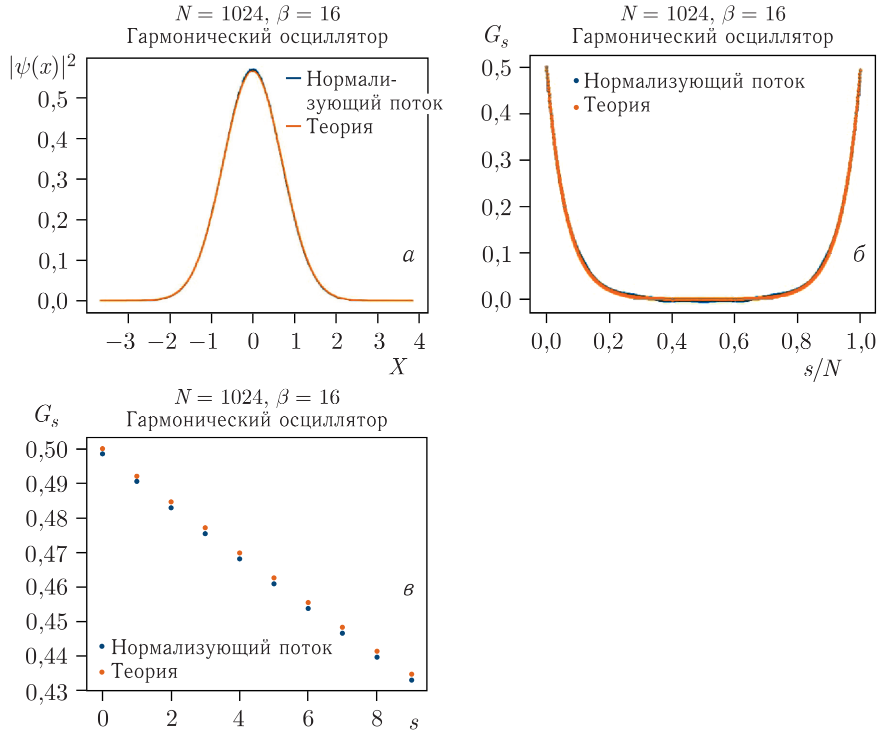

Рис. 1. Результаты вычисления квадрата волновой функции основного состояния (*а*) и двухточечной функции Грина (*б, в*) по траекториям, сгенерированным моделью, в сравнении с теоретическими значениями для гармонического осциллятора

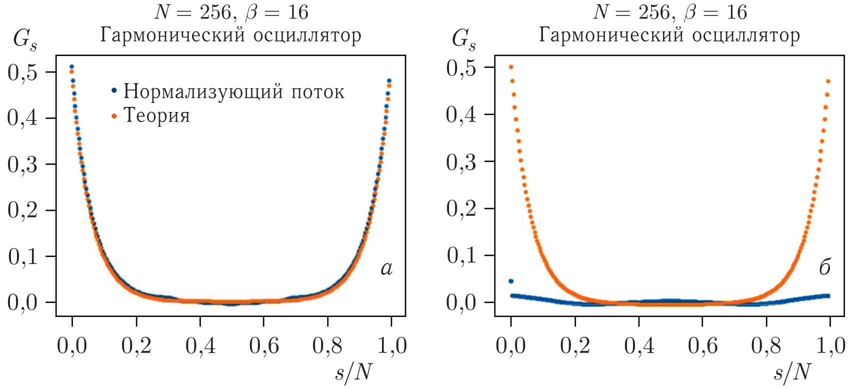

Рис. 2. Сравнение функций Грина, вычисленных с помощью нейросетей, обученных в течение одной эпохи с применением ортогонального преобразования (*а*) и без него (*б*)

4. ДВОЙНАЯ ПОТЕНЦИАЛЬНАЯ ЯМА $H = \frac{p^2}{2} + g\,(x^2 - x_0^2)^2$

Если потенциал $V(x)$ имеет более сложный вид, чем потенциал гармонического осциллятора, то нейронная сеть приближает целевое распределение, вообще говоря, с невысокой точностью. Для точного воспроизведения целевого распределения необходимо траектории, сгенерированные нейросетью, взять в качестве стартовых для алгоритма Метрополиса. При этом алгоритм сойдется быстрее по сравнению с тем случаем, когда в качестве начальных взяты траектории из более простого распределения, например, холодные траектории: $x_i^{(k)} = 0$ (рис. 3).

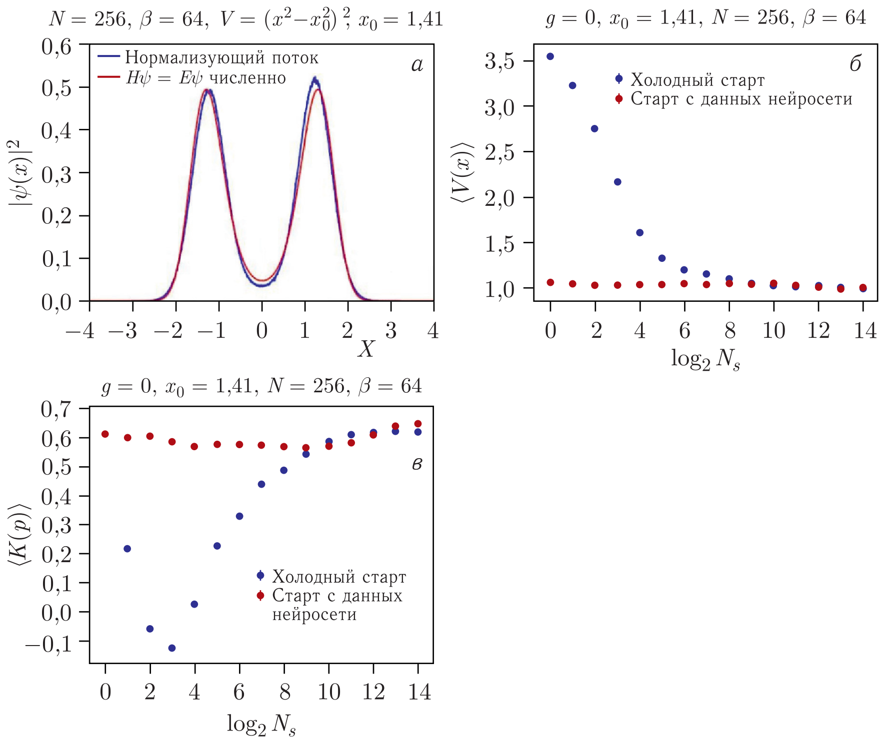

Рис. 3 (цветной в электронной версии). Результаты для модели двойной потенциальной ямы. *а*) Квадрат волновой функции основного состояния, вычисленный по сгенерированным траекториям в сравнении с численным решением уравнения Шрёдингера. *б, в*) Результат применения алгоритма Метрополиса к холодным (синий) и к сгенерированным (красный) траекториям. По горизонтальной оси отложено число итераций алгоритма Метрополиса (свипов) в логарифмическом масштабе, по вертикальной оси отложено среднее значение вычисляемой наблюдаемой: потенциальной энергии (*б*) и кинетической энергии (*в*)

5. ПОТЕНЦИАЛ МОРЗЕ $H = \dfrac{p^2}{2} + \dfrac{1}{2}\left[(\mathrm{e}^{-\alpha x} - 1)^2 - 1\right]$

Волновая функция и средние значения потенциальной и кинетической энергии известны точно [9]:

$$\langle K \rangle = \frac{\alpha}{4}\left(1 - \frac{\alpha}{2}\right), \quad \langle V \rangle = \frac{\alpha}{4} - \frac{1}{2}, \tag{8}$$

$$|\psi|^2 = \frac{2^{2/\alpha - 1}(2-\alpha)}{\alpha^{2/\alpha-1}\Gamma\left(2/\alpha\right)} \exp\left[-\frac{2}{\alpha}\,\mathrm{e}^{-\alpha x} - (2-\alpha)x\right]. \tag{9}$$

Результаты вычислений на решетке (рис. 4, 5) мы сравниваем со значениями (8), (9), которые являются предельными для решеточных величин при $N \to \infty$ (непрерывный предел).

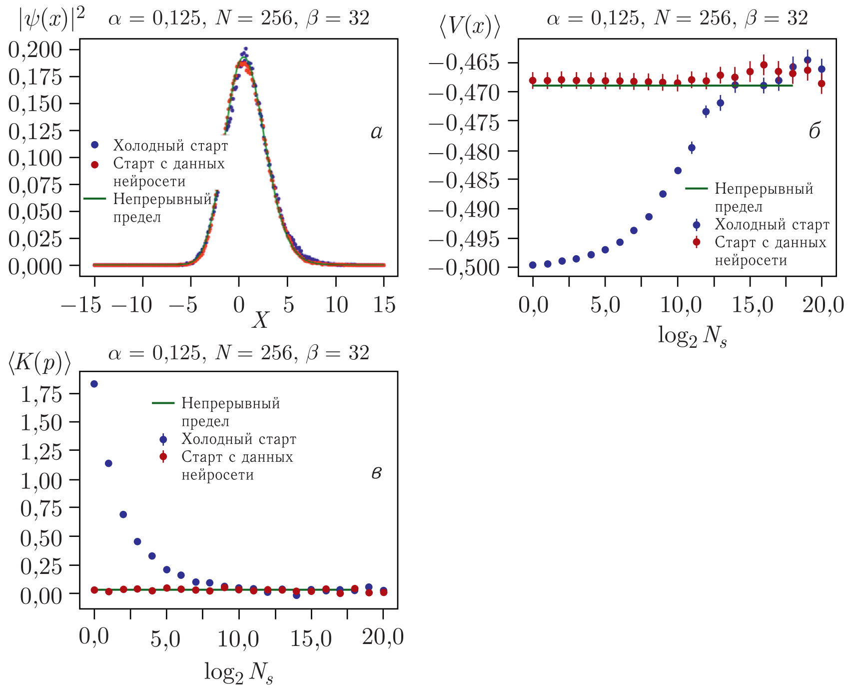

Рис. 4 (цветной в электронной версии). Результаты для модели с потенциалом Морзе при $\alpha = 0{,}125$. *а*) Квадрат волновой функции основного состояния, вычисленный по сгенерированным траекториям в сравнении с аналитическим выражением. *б, в*) Результат применения алгоритма Метрополиса к холодным (синий) и к сгенерированным (красный) траекториям. По горизонтальной оси отложено число итераций алгоритма Метрополиса (свипов) в логарифмическом масштабе, по вертикальной оси отложено среднее значение вычисляемой наблюдаемой: потенциальной энергии (*б*) и кинетической энергии (*в*)

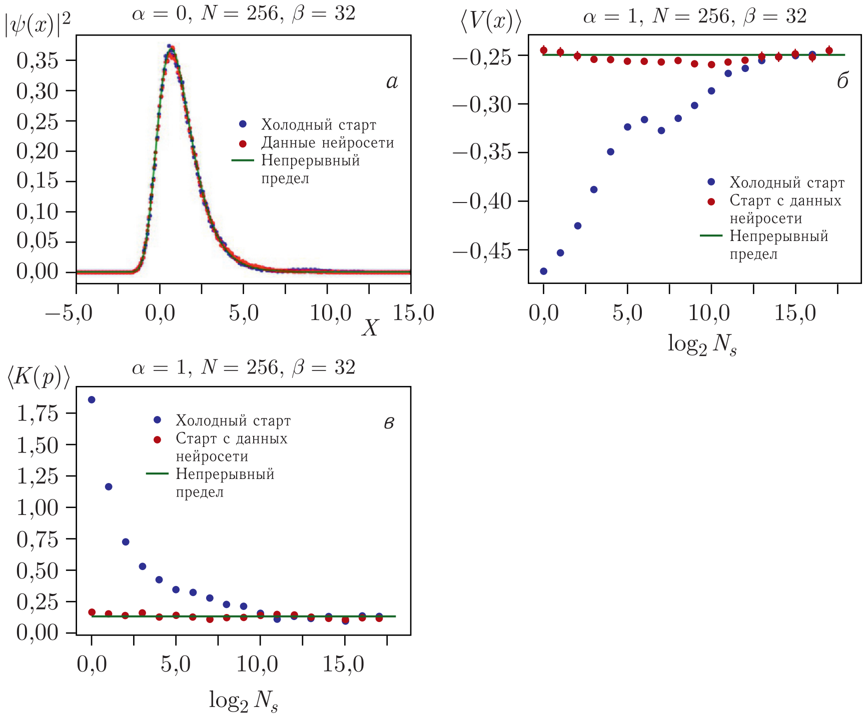

Рис. 5 (цветной в электронной версии). Результаты для модели с потенциалом Морзе при $\alpha = 1$. *а*) Квадрат волновой функции основного состояния, вычисленный по сгенерированным траекториям в сравнении с аналитическим выражением. *б, в*) Результат применения алгоритма Метрополиса к холодным (синий) и к сгенерированным (красный) траекториям. По горизонтальной оси отложено число итераций алгоритма Метрополиса (свипов) в логарифмическом масштабе, по вертикальной оси отложено среднее значение вычисляемой наблюдаемой: потенциальной энергии (*б*) и кинетической энергии (*в*)

ЗАКЛЮЧЕНИЕ

Таким образом, нейросетевой генеративный алгоритм нормализующих потоков позволяет генерировать траектории с распределением, близким к целевому, что многократно ускоряет вычисление средних значений наблюдаемых. В следующих работах предполагается распространить этот подход на модели с более сложным гамильтонианом, а также на $1 + d$-мерную квантовую теорию поля.

Финансирование. Работы Д. В. Сальникова и В. В. Чистякова поддержаны Фондом развития теоретической и математической физики «Базис» (гранты № 23-2-2-40-1 и № 23-2-9-18-1).

СПИСОК ЛИТЕРАТУРЫ

1. *Zinn-Justin J.* Path Integrals in Quantum Mechanics. London: Oxford Univ. Press, 2004.
2. *Creutz M., Freedman B. A.* A Statistical Approach to Quantum Mechanics // Ann. Phys. 1981. V. 132. P. 427–462; https://api.semanticscholar.org/CorpusID: 307835.
3. *Gabrie M., Rotskoff G., Vanden-Eijnden E.* Adaptive Monte Carlo Augmented with Normalizing Flows // Proc. Natl. Acad. Sci. 2022. V. 119, No. 10. P. e2109420119.
4. *Gao C., Isaacson J., Krause C.* i-Flow: High-Dimensional Integration and Sampling with Normalizing Flows // Mach. Learn.: Sci. Technol. 2020. V. 1, No. 4. P. 045023.
5. *Li S. H., Wang L.* Neural Network Renormalization Group // Phys. Rev. Lett. 2018. V. 121. P. 260601; https://link.aps.org/doi/10.1103/PhysRevLett.121. 260601.
6. *Albergo M. S., Boyda D., Hackett D. C., Kanwar G., Cranmer K., Racaniére S., Rezende D. J., Shanahan P. E.* Introduction to Normalizing Flows for Lattice Field Theory. http://arxiv.org/abs/2101.08176. 2021.
7. *Papamakarios G., Nalisnick E., Rezende D. J., Mohamed S., Lakshminarayanan B.* Normalizing Flows for Probabilistic Modeling and Inference // J. Machine Learn. Res. 2021. V. 22, No. 1. P. 2617–2680; arXiv:1912.02762.
8. *Albergo M. S., Kanwar G., Shanahan P. E.* Flow-Based Generative Models for Markov Chain Monte Carlo in Lattice Field Theory // Phys. Rev. D. 2019. V. 100. P. 034515; https://link.aps.org/doi/10.1103/PhysRevD.100.034515.
9. *Landau L. D., Lifshitz E. M.* Quantum Mechanics: Non-Relativistic Theory. V. 3. 3rd ed. Pergamon Press, 1977.